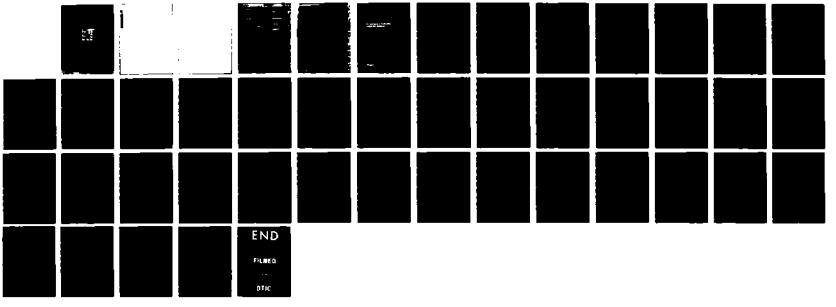
AD-A157 399 THE ANALYSIS OF SUBJECTIVE JUDGMENT MATRICES(U) RAND CORP SANTA MONICA CA G CRAWFORD ET AL. MAY 85 RAND/R-2572-1-AF F49620-82-C-0018 1/1

UNCLASSIFIED F/G 12/1 NL



1.1
1.25

Unclassified

Approved for Public Release; Diatribution Unlimited

No Restrictions

Matrices (Mathematics), Vectors (Mathematics)
Eigenvectors,
Comparison,
Statistical Analysis,

see reverse side

A popular method for quantifying subjective judgment utilizes the dominant eigenvector of a matrix of paired comparisons. The eigenvector yields a scale of the importance of each element of a collection relative to the others. The scale is based on a matrix of subjective paired comparisons. Thomas Saaty has shown this to be a useful tool for analyzing hierarchical structures in many military and industrial applications: by estimating the scale at each level of a structured problem, this procedure yields the relative importance of the elements at the bottom level of the hierarchy to the goals or output at the top level. The geometric mean vector is computationally easier than and statistically preferable to the eigenvector. Further, the geometric mean vector is applicable to a wider class of problems and has the advantage of arising from common statistical and mathematical models. The statistical advantages are theoretically and empirically demonstrated. key words: front of 1473

Accession For	
NTIS GRA&I	X
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

DTIC COPY INSPECTED 3

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

R-2572-1-AF

The Analysis of Subjective Judgment Matrices

Gordon Crawford, Cindy Williams

May 1985

A Project AIR FORCE report
prepared for the
United States Air Force

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

PREFACE

There is a growing literature and interest in methods for quantifying subjective judgments. Several ongoing Air Force efforts utilizing subjective judgment have come to the authors' attention. Mission Area Analysis requires subjective estimates of a large number of parameters. Long-range planning repeatedly draws on judgments about the future importance and worth of plans and geographical areas. The Constant Quest project, directed by the Readiness/NATO Coordination Board, highlighted the importance of subjective judgments in evaluating command and control systems.

Thomas Saaty of the University of Pennsylvania has advanced a popular tool for quantifying and scaling the worth of a set of objects or entities. For problems that fit the Saaty framework, this report details an improvement on Saaty's "eigenvector" technique that is easier to use and more amenable to statistical inferences.

This report was prepared under the Project AIR FORCE research study effort, "Evolving Concepts for Long-Range Planning." It was originally published in 1980 and has been expanded, clarified, and updated in 1985, as part of the Project AIR FORCE Resource Management Program's concept formulation and exploratory research activity.

SUMMARY

Let $\{E_1, E_2, \ldots, E_n\}$ be a collection of objects or entities that are in some sense comparable. The E_i may be alternative plans to achieve some goal, alternative objects that have some comparable utility, or generally a collection of entities that have varying degrees of some common value. A vector $u_1, u_2, \ldots, u_n$ is called a ratio scale for the collection if, for each i and j, u_i/u_j is the ratio of the value of E_i to the value of E_j.

An important application of ratio scales is in the study of hierarchies. Suppose that for each level of a hierarchy there is a ratio scale for the value of objects at that level relative to any object at the next level up. The ratio scales for various levels can be combined multiplicatively to give a view of the entire hierarchy. Because hierarchies are used to model complex systems in many important military and industrial applications, the estimation of the ratio scales deserves considerable attention.

Suppose that a ratio scale $u_1, u_2, \ldots, u_n$, $u_i > 0$, for objects $E_1, E_2, \ldots, E_n$ exists but is not known. Let a_{ij}, $i,j = 1,2, \ldots, n$ be subjective estimates of u_i/u_j made by a judge. In particular, we assume $a_{ii} = 1$ for each i, and $a_{ji} = 1/a_{ij}$. The matrix $A = [a_{ij}]$ of subjective pairwise comparisons is called a judgment matrix.[1]

If the judge is perfectly consistent in making estimates, the matrix A will satisfy the consistency criterion

$$a_{ij}a_{jk} = a_{ik} \text{ for each } i,j,k \ .$$

If this condition is met, any column of the matrix A gives a ratio scale for $\{E_1, E_2, \ldots, E_n\}$. However, judgments are frequently inconsistent, and judgment matrices rarely satisfy the consistency criterion. A mathematical procedure is required for estimating an underlying ratio scale based on an inconsistent judgment matrix A.

Thomas Saaty (1977a–d) argues that the "dominant" right eigenvector corresponding to the maximal eigenvalue should be used to estimate the underlying scale. The argument is: The dominant eigenvector is a continuous function of the elements of the matrix, and, if the matrix is consistent, the eigenvector gives the unique (to within scalar multiplication) scale. Thus, if the elements of the matrix get perturbed slightly in the process of being subjectively quantified by a judge, the dominant eigenvector will return a scale only slightly different from the scale of an underlying consistent judgment matrix.

Although the classical analyst may worry about uniform continuity or other erudite intricacies of this argument, we are worried about a more basic oversight: The eigenvector is not the only continuous vector-valued function of judgment matrices that yields the correct scale when the matrix happens to be consistent. There are many others, including the vector of row sums, the vector of the inverse of column sums, any column of the matrix, and the whole ring generated by positive linear combinations of these and other solutions.

We are aware of the desirable properties of the eigenvector in characterizing a linear operator and its spectral decomposition, but the immediate relevance of these properties to this estimation problem seems open to question. In most estimation problems, the wealth of statistical literature on estimation procedures and their properties has enhanced understanding of

[1] In practice, a judge is usually asked to supply the $n(n-1)/2$ upper off-diagonal terms. The remainder of the matrix is defined by the above relationship.

the problem. Below we relate this particular problem to well-known statistical models. The geometric mean vector[2] $v = v_1, v_2, \ldots, v_n$, given by

$$v_i = \prod_{j=1}^{n} a_{ij}^{1/n} ,$$

which satisfies the continuity and consistency criteria Saaty uses to defend the dominant eigenvector, has several other desirable traits: In certain circumstances, it is statistically optimal and gives rise to an estimate of scales and a measure of consistency that have known statistical distributions. In empirical studies reported here it seems to do as well as, or better than, the eigenvector in preserving rank order. In addition, it is supported by a literature describing methods of handling a wealth of variations of the problem, including missing data and multiple judges.

---

[2]Elsewhere in the literature this procedure, and this vector, have been referred to as the Logarithmic Least Squares Method (LLSM).

ACKNOWLEDGMENTS

This work has benefited from the interest and help of several Rand colleagues. Stephen Drezner first questioned the eigenvector approach. Emmett Keeler and Daniel Relles suggested other approaches that helped formulate the problem. The late Edwin Paxson used the procedure we advocate with little more justification than his strong intuition.

Gus Haggstrom's keen interest and insights kept the problem alive and helped place it in perspective with other statistical research. Captain Jordan Kriendler and Major Michael Parmentier of AF/PAXE have kindly let us use their example in several sections of this report.

Our thanks to Helen Turin for trying to make it readable.

CONTENTS

PREFACE iii

SUMMARY v

ACKNOWLEDGMENTS vii

FIGURE AND TABLES xi

Section

I. INTRODUCTION 1
- Quantitative Analysis of Subjective Data 1
- Eliciting Subjective Judgments 2
- Synthesizing Data in Large Problems 3
- Constructing Measurement Scales 4

II. PAIRWISE COMPARISONS, THE JUDGMENT MATRIX, AND THE ESTIMATION PROBLEM 5

III. CONSISTENT MATRICES AND RATIO SCALES 7

IV. SAATY'S NORMALIZED EIGENVECTOR SCALE 9

V. APPLICATIONS OF THE RATIO SCALE 11

VI. THE GEOMETRIC MEAN SCALE 14

VII. THE GEOMETRIC MEAN VECTOR AND THE MAXIMUM LIKELIHOOD ESTIMATOR 18

VIII. EMPIRICAL COMPARISON OF THE TWO METHODS: A MONTE CARLO STUDY 20

IX. EXAMPLE 26

Appendix: COMPARISON OF RANK PRESERVATION ABILITIES 29

BIBLIOGRAPHY 33

FIGURE

1. Hierarchy for selecting an automobile 12

TABLES

1. Ratio Scales 13
2. Ratio Scales Perturbed by Lognormal Errors, Sum of Squared Errors of Logarithms, Monte Carlo Comparison of Geometric Mean Vector and Eigenvector . 22
3. Ratio Scales Perturbed by Ratios of Uniform Random Variables, Sum of Squared Errors of Logarithms, Monte Carlo Comparison of Geometric Mean Vector and Eigenvector 23
4. Ratio Scales Perturbed by Lognormal Errors, Sum of Squared Errors, Monte Carlo Comparison of Geometric Mean Vector and Eigenvector 24
5. Ratio Scales Perturbed by Ratios of Uniform Random Variables, Sum of Squared Error, Monte Carlo Comparison of Geometric Mean Vector and Eigenvector 25
6. Comparison of Scale Estimates for the Automobile Example 27
7. Determination of Rankings of Five Automobiles 28
A.1. Ratio Scales Perturbed by Lognormal Errors, Rank Preservation Error, Monte Carlo Comparison of Geometric Mean Vector and Eigenvector 30
A.2. Ratio Scales Perturbed by Ratios of Uniform Random Variables, Rank Preservation Error, Monte Carlo Comparison of Geometric Mean Vector and Eigenvector 31

I. INTRODUCTION

Over the past three decades, psychometric, military, and industrial researchers have directed considerable effort to the quantitative analysis of subjective data. Analytic tools building upon the subjective judgments of experts have been used in such diverse fields as energy policy analysis, marketing research, economic forecasting, and military planning. Problems amenable to the analysis of subjective information abound, and numerous methods have been proposed for acquiring and treating judgmental data.

One question that arises in the treatment of subjective data is how to construct a scale of relative merit for a collection of objects or activities based upon subjective comparisons of each pair in the collection. For example, consider a collection of three objects, labeled A, B, and C. Suppose that an expert believes A has twice the merit of B, B has three times the merit of C, and A has six times the merit of C. It is natural to construct a scale of relative merit for A, B, and C as (1, 1/2, 1/6). However, suppose that the expert says A has twice the merit of B and B has three times the merit of C, but A has only four times the merit of C. In this case, it is not so easy to decide upon a scale for A, B, and C. This sort of inconsistency is common in human judgments, especially when complicated issues are involved.

Thomas Saaty of the University of Pennsylvania has developed a matrix eigenvector procedure for constructing scales of merit based on inconsistent pairwise comparisons. The method has been applied in a wide variety of planning and decision problems.

This report presents an alternative approach that is preferable to the eigenvector procedure in several important respects. The proposed procedure is derived within a statistical framework and is compared with the eigenvector method on the basis of theoretical and empirical considerations.

The remainder of this introduction discusses the motivation for dealing formally and quantitatively with subjective information and provides a brief review of some of the literature. Section II provides a short, nonrigorous discussion of the eigenvector method and the proposed method for utilizing subjective judgments in quantitative analysis. An example, illustrating the use of the two methods as well as similarities and differences in their results, is introduced in this section and examined throughout the report.

In Section III we give rigorous definitions and develop a framework for treating the estimation problem with classical statistical techniques. Section IV provides a mathematical treatment of the eigenvector method. Section V deals with the application of subjective judgment methods to the study of hierarchical structures. The example introduced in Section II is considered in further detail there.

Section VI introduces the geometric mean vector and gives theoretical justification for its use as an estimator of subjective scales. In Section VII we define a statistical measure of consistency for subjective judgment matrices. Section VIII presents results of a Monte Carlo study comparing the two methods. Section IX considers in greater detail the example introduced in Section II and expanded on in Section V.

QUANTITATIVE ANALYSIS OF SUBJECTIVE DATA

The quantification of subjective data is essential for dealing with a wide class of problems whose solution by other methods would be extremely difficult or impossible. Such problems

are often amorphous and vaguely stated. They involve large, multifaceted issues of importance to decisionmakers and interest groups with diverse backgrounds and biases. Their outcomes may determine the allocation of large sums of public money and impinge critically on the public interest. Moreover, some facets of the problems may lack any well-defined, scalar-valued measures of merit. Even if there are appropriate measures, the collection of relevant objective data might be prohibitively expensive or impossible.

Such problems frequently arise in the assessment of future needs for large organizations. As an example, consider the problem of long-range planning in the U.S. Air Force. This problem involves a great many interrelated issues: the effects of political and economic factors on national security, the importance of various geographic regions of the world to U.S. interests, the threat posed by conflicts of different types in different regions, the current strength of forces to deal with such conflicts, and so on. Although it might be possible to define objective yardsticks to deal with some of these issues, it certainly is not possible for all of them. For some issues, subjective judgments of relative importance or value are the only measures available.

In some problems the best information available is subjective, so why is quantitative analysis desirable at all? Why not just ask the experts to make plans and decisions based on an informal, intuitive analysis? In fact, problems that are not amenable to hard analysis are frequently resolved through intuitive analysis by experts and decisionmakers. However, there are several good reasons for using a formal, quantitative approach in these problems.

A formal analytic framework gives structure and definition to an amorphous mass of data. It allows the decisionmaker to consider relevant information systematically and to examine options and consequences one at a time. In such a framework, the analyst can break an unmanageable problem into manageable parts and then synthesize information about the parts in a rational fashion.

A formal analytic framework also permits sensitivity analysis on alternative judgments. When a problem is considered within a formal framework, tradeoffs among alternative judgments can be spelled out explicitly, and the effects of variations in subjective judgments on outcomes can be studied. Sensitivity analysis may even provide a basis for resolving different points of view.

Perhaps the greatest advantage of a formal analysis, especially in governmental policymaking, is that it is repeatable. Formal analysis provides the audit trial that is so important in matters involving extensive allocation of public resources and impinging on the public interest.

Research literature on the use of subjective information emphasizes three major issues: how to elicit meaningful subjective judgments from individuals or groups, how to synthesize subjective and objective data obtained from various facets of a large problem, and how to construct measurement scales based on subjective information. Following is a review of some of the literature related to each of these issues.

ELICITING SUBJECTIVE JUDGMENTS

Methods for eliciting subjective judgments have received considerable attention in operations research and forecasting literature. Two such methods are war gaming and scenario writing, both of which are used extensively in military planning to provide insights into possible future environments and needs.

Much of the literature on eliciting judgments deals with the problem of acquiring a collective expert opinion free from the usual negative effects of group pressure. An important

method in this category is the Delphi technique, a controlled feedback procedure originally developed by researchers at The Rand Corporation (Gordon and Helmer, 1964). In Delphi, a researcher interrogates a group of experts individually concerning their opinions on possible future events. The researcher assembles means and quartiles for quantitative data thus obtained and presents them individually to group members along with arguments and comments made by other members. Group members can then revise their judgments. The procedure is repeated until the range of judgments narrows. The controlled feedback mechanism in Delphi makes it possible for a group of experts to avoid the usual social pressures of open discussion. The method has been used in many military and industrial applications (see, e.g., Ayres, 1969; Linstone and Turoff, 1975).

The Delphi technique has given rise to several modifications. The Probe method designed by researchers at TRW for forecasting technological events combines Delphi with a timing chart structure so that events can be considered in sequence (North and Pyke, 1968). The method of qualitative controlled feedback proposed by Press (1979) is similar to Delphi in that it uses a controlled feedback loop to aid groups in arriving at judgments, but it differs in that at each iteration, members are supplied only arguments and comments from the group, with no information about the quantitative distribution of group answers.

SYNTHESIZING DATA IN LARGE PROBLEMS

Economists and statisticians have proposed various methods for breaking large problems into smaller pieces and quantitatively synthesizing subjective and objective data from the pieces. One of the most popular of these methods is multi-attribute utility theory, which provides a framework for selecting an optimal decision from among multiple alternatives when some effects of the decision can be measured only subjectively. The expected value of each alternative is determined as a function of the decisionmaker's preferences for the possible consequences and the probabilities that the alternative will lead to those consequences. The probabilities are generally determined from subjective judgments. Some of the decisionmaker's preferences are determined on the basis of subjective indexes such as aesthetic appeal, and others are determined on the basis of objective measures such as cost. The alternative with maximum expected value is chosen as the optimum decision.

The mathematical foundation for multi-attribute utility theory was laid by von Neumann and Morgenstern (1947). Application of the theory to business problems was pioneered by Raiffa and extended by Keeney and others. The theory has been applied to many problems in industrial, government, and military settings (see Keeney, 1973; Keeney and Nair, 1974; Kelley, 1976; MacCrimmon, 1969; Raiffa, 1969). A book by Keeney and Raiffa (1976) gives an excellent treatment of the subject.

A similar method was applied to military problems in a 1958 Master's thesis by Wells, who gives a detailed framework for assessing the relative desirability of existing or proposed weapon systems. System desirability is determined as a function of feasibility, cost, and an attribute Wells calls "military worth." Wherever possible, objective measures are used to evaluate these three factors, and expert judgments are used where there are no objective measures. In particular, military worth is an aggregate property evaluated by analyzing a complex hierarchy and subjective scales for several variables. The Honeywell Corporation used Wells's method in a military planning model called PATTERN (Sigford and Parvin, 1965). A detailed description of the method can be found in Wells (1967).

Saaty proposes that complex decision problems be viewed in terms of hierarchies of objects or properties. At each level of a hierarchy, Saaty uses subjective judgments to estimate a merit scale of the objects. Scales from all the levels are combined mathematically to provide quantitative information about the whole problem. The results in this report are applied within the framework of hierarchical analysis proposed by Saaty. Hierarchies are discussed in Section V.

CONSTRUCTING MEASUREMENT SCALES

Many methods have been developed for constructing scales of measurement based on subjective data. Several books and hundreds of articles have been written about these methods. A classic reference for early contributions, especially for work on psychophysical scales, is Torgerson's (1958) book.

Churchman and Ackoff (1954) did pioneering work in the area of estimating scales of values for decision problems. They used a criterion of additive order consistency to estimate scales from successive subjective judgments. Their paper described several applications to industrial problems. Wells and others later applied the Churchman and Ackoff method in military decision problems.

Much of the literature on subjective scales concerns the estimation of scales from pairwise comparison data. Much statistical work in this area goes under the name "paired comparisons." In the simplest paired comparison experiment, each of several judges examines a number of objects two at a time and states which of the two objects is preferred. No indication of strength of preference is given. Data from these paired comparisons are then used in a statistical model to estimate a scale of preference for the objects. Such an experiment might be used by marketing researchers to determine the relative taste appeals of several new food items.

A good reference for the statistical theory of paired comparisons is David's (1963) book. A bibliography of recent articles on the subject was compiled by Davidson and Farquhar (1976).

Saaty has proposed another method for estimating subjective scales using pairwise comparisons in which a single judge makes pairwise comparisons of a number of objects. For each pair, the judge states not only which object is preferred, but to what degree that object is preferred over the other. A preference scale is determined for the objects based on an eigenvector analysis of the matrix of pairwise comparisons.

Saaty has published many articles (see the bibliography) describing the eigenvector procedure for estimating subjective scales and illustrating the usefulness of this procedure in analyzing complex hierarchical structures. He has applied the procedure in a broad range of problems in the social sciences (Saaty, 1977b; Saaty and Bennett, 1977; Alexander and Saaty, 1977). The procedure has also gained acceptance in military applications and is currently being used as a tool in Air Force long-range planning.

If one assumes that an analysis problem is in the context that Saaty has been writing about, then the results of this report imply that the geometric mean is preferable to the eigenvector solution. However, Veit and Callero (1981) and Veit, Callero, and Rose (1982, 1984) believe that many problems may not be multiplicative at all and may not fit the Saaty context.

II. PAIRWISE COMPARISONS, THE JUDGMENT MATRIX, AND THE ESTIMATION PROBLEM

Consider the problem of purchasing a new car.[1] Suppose that a preliminary investigation yields five specific makes that seem appropriate. The price of each make is known, and although some other measures of merit may have been quantified (principally performance measures), the important subjective question of how much each car satisfies the overall needs is difficult to quantify.

We will attempt to assign to each make of automobile an estimate of utility in such a way that if u_i is the utility of the ith make, then u_i/u_j is a measure of the preference of the ith make to the jth make. The vector $u_1, u_2, \ldots, u_5$ will be called a *ratio scale*.

Some aspects of the usefulness of such a ratio scale are immediately apparent. We could, in this example, choose between the cars on the basis of utility per dollar of initial cost or, with more foresight, on the basis of utility per dollar of expected life cycle cost.

To estimate the vector of utilities $u_1, u_2, \ldots, u_5$ Saaty has suggested the following procedure (see especially 1977b): We construct a matrix composed of our subjective estimates of the ratios of the utilities of all possible pairwise combinations, so that the elements a_{ij} of the matrix A are our estimates of u_i/u_j. Thus we know that the diagonal elements are given by $a_{ii} = 1, i = 1, \ldots, 5$. Additionally, the lower off-diagonal elements are determined by the upper off-diagonal elements: $a_{ji} = 1/a_{ij}$.

Saaty (1977b) proves that in this case the matrix A has a maximal eigenvalue and a corresponding eigenvector (the dominant eigenvector) all of whose components are positive. Saaty proposes, primarily with empirical justification, that this dominant eigenvector be used as an estimate of the ratio scale.

Suppose that when we form our estimates, the relative utility of Make 1 to Make 2 is 2, of Make 1 to Make 3 is 1/9, of Make 1 to Make 4 is 1/6 and of Make 1 to Make 5 is 1/7. Then the first row of our judgment matrix has the form

$$1, \quad 2, \quad 1/9, \quad 1/6, \quad 1/7 \ .$$

Continuing, suppose that we have filled in the upper off-diagonal of our judgment matrix:

$$\begin{array}{ccccc} 1 & 2 & 1/9 & 1/6 & 1/7 \\ & 1 & 1/9 & 1/6 & 1/7 \\ & & 1 & 6 & 4 \\ & & & 1 & 1/7 \\ & & & & 1 \end{array}$$

Then, in view of reciprocal symmetry we have

$$A = \begin{array}{ccccc} 1 & 2 & 1/9 & 1/6 & 1/7 \\ 1/2 & 1 & 1/9 & 1/6 & 1/7 \\ 9 & 9 & 1 & 6 & 4 \\ 6 & 6 & 1/6 & 1 & 1/7 \\ 7 & 7 & 1/4 & 7 & 1 \end{array}$$

[1] This example, due to Capt. Jordan Kreindler and Major Michael Parmentier of the U.S. Air Force Directorate for Programs Evaluation, Systems Analysis Division, is treated in further detail in Sections V and IX.

Continuing with this example we compute the dominant eigenvector w of this matrix and get:

$$w = \begin{matrix} .0378 \\ .0294 \\ .5239 \\ .1131 \\ .2958 \end{matrix}$$

Thus, in this case, our estimate of the utility of the first make is .0378 and of the third make is .5239. For a detailed treatment of this procedure see Saaty, 1977b; Saaty and Bennett, 1977.

This example is discussed in more detail in Sections V and IX, where it is expanded to illustrate the value of ratio scales in analyzing hierarchical structures. In application to hierarchies it is assumed that the objects at each level of the hierarchy depend on the objects of the next lower level in some way. The procedure enables the user to estimate the influence each object in a level has on all the objects or goals in superior levels.

For problems where the eigenvector procedure is useful there is another estimation procedure that is preferable in several respects.

Where Saaty would estimate the utility of the ith object with the ith component of the dominant eigenvector, we argue that a better estimator is given by the vector $v = v_1, v_2, \ldots, v_n$, where

$$v_i = \prod_{j=1}^{n} a_{ij}^{1/n}$$

is the geometric mean of the elements in the ith row of A. In the example above this yields the following estimates:

Object	Eigenvector	Geometric Mean Vector
1	.0378	.0409
2	.0294	.0310
3	.5239	.5307
4	.1131	.1132
5	.2958	.2842

Compared with the dominant eigenvector, the geometric mean vector

1. Is statistically better;
2. Is easier to calculate;
3. Gives rise to a more meaningful measure of consistency that has known statistical properties, allowing tests of hypotheses, confidence interval estimation, etc.;
4. Gives rise to estimates of utility with known statistical properties, allowing tests of hypotheses, confidence interval estimation, etc.;
5. Is supported by statistical literature discribing methods of handling a wealth of variations of the problem;
6. Is rooted in a mathematical approach to estimation that provides an intuitive understanding of the problem and a means for assessing suitability of the method.

III. CONSISTENT MATRICES AND RATIO SCALES

Consider a set of n activities or objects $E_1, E_2, \ldots, E_n$, which contribute to some objective. Suppose the activities can be ranked on a ratio scale $(u_1, u_2, \ldots, u_n)$, $u_i > 0$, so that u_i/u_j measures the degree to which E_i is more important than E_j in achieving the objective. In particular, $u_i/u_j > 1$ if E_i is more important than E_j. Let $A = [a_{ij}]$ be the $n \times n$ matrix of pairwise comparisons of $E_1, E_2, \ldots, E_n$ given by

$$a_{ij} = \frac{u_i}{u_j} \quad i,j = 1,2,\ldots,n \ . \tag{3.1}$$

Then A has the property that

$$a_{ji} = \frac{1}{a_{ij}} \quad i,j = 1,2,\ldots,n \ , \tag{3.2}$$

and in particular

$$a_{ii} = 1, \quad i = 1,2,\ldots,n \ .$$

A square matrix A with positive entries satisfying (3.2) will be called a *judgment matrix*.

It follows immediately from (3.1) that

$$a_{ij}a_{jk} = a_{ik} \ . \tag{3.3}$$

A matrix with positive entries satisfying (3.3) is said to be *consistent*. It is easy to see that every consistent matrix is a *judgment matrix*.

Let A be an arbitrary consistent matrix. Because

$$a_{jk} = \frac{a_{1k}}{a_{1j}} \quad \text{for any } jk \ ,$$

every element of A can be determined from the first row of A. It follows that A is a matrix of rank one with exactly one nonzero eigenvalue. Moreover, it follows from (3.3) that

$$A^2 = nA \ .$$

Thus any column of A is an eigenvector of A, and the single nonzero eigenvalue of A is n.

Let $w = w_1, w_2, \ldots, w_n$ be any eigenvector corresponding to the eigenvalue n. For any k, the kth column of A is an eigenvector corresponding to the same eigenvalue; therefore for each i and j,

$$w_i = ca_{ik}$$

$$w_j = ca_{jk}$$

for some $c \neq 0$, and therefore

$$\frac{w_i}{w_j} = a_{ij} \ .$$

Thus, w is a ratio scale for A. In fact, it is clear that there are infinitely many such scales, each one corresponding to a different scalar multiple of the kth column of A.

The normalized eigenvector with components

$$\frac{w_i}{\sum_{i=1}^{n} w_i}$$

is the particular scale that Saaty (1977b) associates with the consistent matrix A.

A ratio scale corresponding to a consistent matrix A can be derived in several ways. Any column of A is such a scale. The vector of reciprocals of elements of an arbitrary row of A is also a ratio scale for A. It is easy to see that the vector r of row sums defined by

$$r_i = \sum_{j=1}^{n} a_{ij}$$

and the geometric mean vector v defined by

$$v_i = \prod_{j=1}^{n} a_{ij}^{1/n}$$

also provide ratio scales for A. When these scales are normalized, they are equal to the normalized eigenvector scale for a consistent matrix.

IV. SAATY'S NORMALIZED EIGENVECTOR SCALE

Consider again the activities $E_1, E_2, \ldots, E_n$ that contribute to some objective. Suppose a judge makes pairwise comparisons on some scale of the relative importance of each pair of activities with respect to the underlying objective. If a_{ij} represents the relative importance of E_i over E_j, so that $a_{ij} > 1$ if and only if E_i is more important than E_j, it is then natural to insist that the judge make comparisons in such a way that

$$a_{ij} = \frac{1}{a_{ji}} \quad \text{for each } i,j.$$

In other words, such a pairwise comparison matrix is a judgment matrix. The ideal pairwise comparison matrix would also be consistent. For example, if E_i is twice as important as E_j and E_j is three times as important as E_k, one would expect E_i to be six times as important as E_k. However, human judgment is often inconsistent, and it is not likely that a judge making pairwise comparisons will construct a consistent matrix except in cases where the dimension is small. A simple example in which pairwise comparisons do not result in a consistent matrix is that of a tournament: X may win against Y and Y against Z, but X may lose to Z.

The problem we consider is this: Given an inconsistent judgment matrix A, how can we construct a ratio scale that in some sense best reflects the information in the matrix? Saaty (1977b) proposes that the appropriate scale is the normalized eigenvector corresponding to the maximal eigenvalue of A.

Saaty argues as follows: If the judgment matrix A is consistent, then the normalized eigenvector corresponding to the single nonzero eigenvalue n does give the underlying ratio scale. A theorem of Frobenius for matrices with positive entries (Franklin, 1968) guarantees that any judgment matrix has a positive eigenvalue L that exceeds all the other eigenvalues in absolute value. This maximal eigenvalue has an associated eigenvector that is positive in all its components.

Now an inconsistent judgment matrix can be viewed as having been derived from a consistent one by perturbation of some or all of the matrix components. Because the eigenvalues and eigenvectors of a matrix depend continuously on its components, small perturbations in the components will result in small changes in the eigenvalues and eigenvectors. Thus when the perturbations of the components are small, the maximal eigenvalue is close to n, and the corresponding normalized eigenvector is close to the normalized eigenvector of the unperturbed consistent matrix. Therefore, Saaty selects the suitably normalized eigenvector associated with the maximal eigenvalue as the ratio scale corresponding to the judgment matrix.

Saaty also proposes an index of consistency for judgment matrices. He shows that an $n \times n$ judgment matrix whose only nonzero eigenvalue is n must be consistent, and that the maximal eigenvalue L for an inconsistent judgment matrix is strictly greater than n. Therefore he uses the normalized difference

$$\mu = \frac{L-n}{n-1}$$

as the index of consistency of an $n \times n$ judgment matrix with maximal eigenvalue L. Notice that the index is zero for consistent matrices and positive for inconsistent ones. Saaty also

shows that the index increases as perturbations of the components away from consistency increase.

Unfortunately, the question of how small the perturbations of matrix components must be to give rise to a given deviation in the maximal eigenvalue is a delicate one. Saaty (1977b) describes an empirical investigation of this question in which he determines the consistency indices corresponding to randomly generated judgment matrices of different dimensions. However, because the eigenvector does not fit into any standard statistical framework, there is no readily available device against which deviations from consistency can be measured.

Saaty does show (1977b) that the consistency index μ reflects the variance in judgmental errors for an inconsistent matrix in the following sense: Suppose that the pairwise comparisons a_{ij} in the judgment matrix A actually arise from perturbations of the ratios of components of some underlying scale $u_1, u_2, \ldots, u_n$; i.e.,

$$a_{ij} = \frac{u_i}{u_j} e_{ij},$$

$$e_{ij} = 1 + d_{ij}.$$

Saaty shows that for small d_{ij}, 2μ is an estimate of the variance of the d_{ij}. Starting from this estimate, Saaty (1977b) develops a test of the hypothesis of consistency for a judgment matrix.

The choice of a scale to be used in filling in a pairwise comparison matrix is somewhat arbitrary. Because people find it difficult to rank more than about seven objects at a time, Saaty recommends a subjective pairwise comparison scale consisting of the integers from one to nine together with their reciprocals. In this scale, a value of 1 is assigned to pairs of objects that are equally important. The integers 3, 5, 7, and 9 are associated with descriptive words (9 means "absolute importance," 5 means "essential or strong importance"), and the integers 2, 4, 6, and 8 are used for intermediate values. Reciprocals of integers are used so that the matrix of pairwise comparisons is a judgment matrix—that is, is reciprocal symmetric.

V. APPLICATIONS OF THE RATIO SCALE

Saaty presents numerous applications requiring the estimation of ratio scales from pairwise comparison information (see the bibliography). He cites examples in economics, political science, and transportation planning, as well as in personal planning areas such as choosing a school or a vacation spot.

One of the most interesting and useful applications of the ratio scale is in the study of hierarchical systems. A hierarchy is a collection of objects grouped according to levels. Objects at a given level of the hierarchy depend on objects at lower levels. The objects at one level may be ranked on a ratio scale according to their importance relative to a given object at the next higher level. Thus one may construct a system of ratio scales, one scale for each level relative to every object in the next level up.

Once such a system of ratio scales has been constructed, it can be used to study interactions among all levels of the hierarchy. For example, in a hierarchy consisting of three levels, we may determine the ranked importance of objects on the lowest level relative to each object in the highest level. Suppose for simplicity that the highest level consists of a single object, that the second level has n objects, and the third has m objects. Let $(w_1, w_2, \ldots, w_n)$ be the ratio scale that reflects the importance of objects in the second level relative to the single object in the first level. Now the objects in the third (lowest) level may be ranked on a ratio scale relative to each object on the second level. Let

$$u_{1j}, u_{2j}, \ldots, u_{mj}, \quad j = 1,2, \ldots, n$$

be the ratio scale for the third level relative to the jth object of the second level. Then, according to Saaty, the importance of objects in the lowest level relative to the highest level may be measured by the vector $v_1, v_2, \ldots, v_m$, where

$$v_i = \sum_{j=1}^{n} u_{ij} w_j, \quad i = 1,2, \ldots, m .$$

In matrix notation, if $w = w_1, w_2, \ldots, w_n$ is the ratio scale for the second level relative to the single object in the first level, and if

$$U = \begin{matrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ u_{m1} & u_{m2} & \cdots & u_{mn} \end{matrix}$$

is the matrix whose jth column is the ratio scale for the objects in the third level relative to the jth object in the second level, then

$$v = Uw$$

gives a scale of importance of objects in the third level relative to the first.

The same procedure may be extended to hierarchies with more than three levels. Thus knowing only the measures of importance of objects in each level relative to individual objects in the adjacent higher level, we may deduce their ranked importance relative to objects at all higher levels. In particular, objects at the lowest level can be ranked according to their importance relative to the object (or objects) at the highest level.

As an example, consider the problem of selecting an automobile. (This example is treated in more detail in Section IX.) The problem can be viewed in terms of the hierarchical structure shown in Fig. 1. The highest level of the hierarchy is the final selection of the automobile. On the second level are attributes of the automobiles—namely status, cost, economy, and size. The third level consists of the automobile makes to be considered. The automobiles themselves are ranked according to each of the attributes, and the attributes are ranked according to their importance relative to the overall objective of selecting a car.

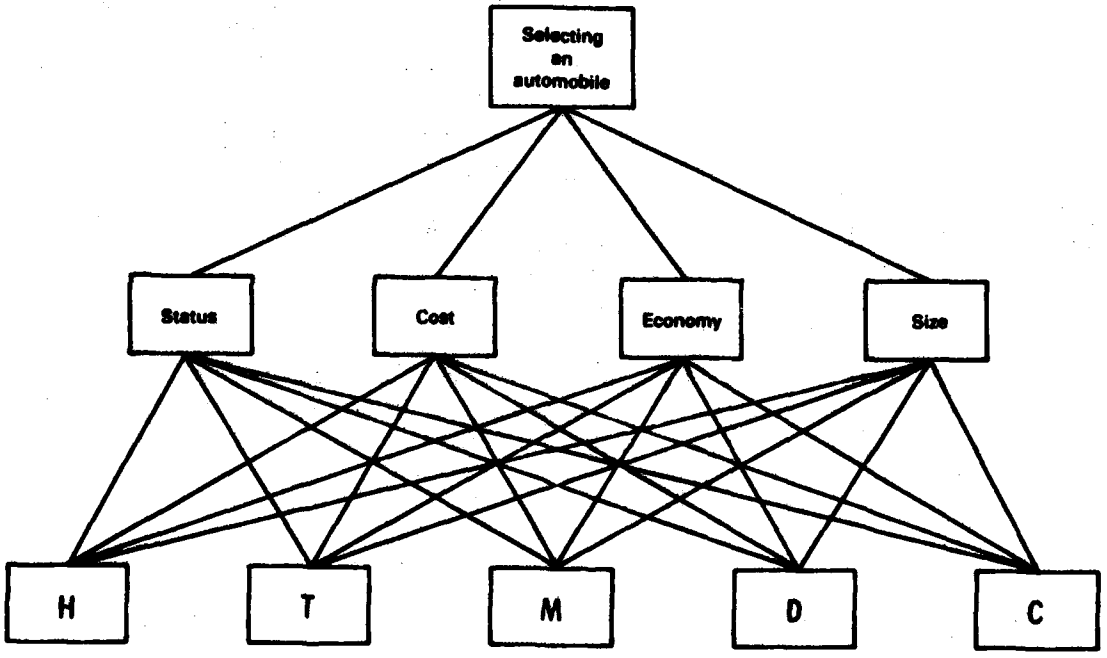


Fig. 1—Hierarchy for selecting an automobile

Table 1 gives the ratio scales determined from judgments made by one prospective buyer. The order of automobile preference for this buyer is then given by the product of the matrix and the vector in Table 1 as follows:

$$\begin{matrix} .0378 & .2703 & .3357 & .3143 \\ .0294 & .4196 & .2929 & .4630 \\ .5239 & .0554 & .1071 & .0596 \\ .1131 & .1571 & .1500 & .1288 \\ .2958 & .0976 & .1143 & .0343 \end{matrix} \cdot \begin{matrix} .0885 \\ .2425 \\ .2579 \\ .4112 \end{matrix} = \begin{matrix} .2847\ H \\ .3703\ T \\ .1119\ M \\ .1397\ D \\ .0934\ C \end{matrix}$$

The final ranking reflects the buyer's perception of the relative status, cost, economy, and size of the five automobiles considered as well as his judgment of the relative importance of these four attributes in the selection of an automobile. This buyer's first choice should be T and his last choice C.

Table 1

RATIO SCALES

Ratio Scale of Second Level Relative to First Level	
Status	.0885
Cost	.2425
Economy	.2579
Size	.4112

Ratio Scales of Third Level Relative to Attributes at the Second Level				
	Status	Cost	Economy	Size
H	.0378	.2703	.3357	.3143
T	.0294	.4196	.2929	.4630
M	.5239	.0554	.1071	.0596
D	.1131	.1571	.1500	.1288
C	.2958	.0976	.1143	.0343

VI. THE GEOMETRIC MEAN SCALE

Saaty's examples clearly show that the study of interactions among various levels of a hierarchy depends heavily upon an assessment of the ranked importance of objects at each level relative to objects in the level above. The basic building blocks in a hierarchical study are the ratio scales measuring the relative importance of objects at a given level. One would like assurance that the estimates of the ratio scales are well grounded in statistical theory and that they work well empirically. Below we propose a method for constructing a ratio scale based on pairwise comparisons. That method seems superior to the eigenvector procedure when judged by the criteria mentioned above.

For $n \times n$ judgment matrices $A = [a_{ij}]$ and $C = [c_{ij}]$, define

$$m(A,C) = \left[\sum_{j=1}^{n} \sum_{j>1} (\ell n\ a_{ij} - \ell n\ c_{ij})^2 \right]^{1/2} .$$

It is not difficult to verify that m satisfies the triangle inequality and is a metric on the space of $n \times n$ judgment matrices. *Theorem 3* will show that for any $n \times n$ judgment matrix A, there is a consistent matrix C that is m-closest to A. Such a consistent matrix is given by

$$c_{ij} = \frac{v_i}{v_j} ,$$

where

$$v_i = \prod_{j=1}^{n} a_{ij}^{1/n} ;$$

that is, v_i is the geometric mean of the elements of the ith row of A. We will use the vector v, suitably normalized, as the estimate of the ratio scale corresponding to A.

The following two invariance properties show that m is a suitable choice of metric for the space of judgment matrices. Their proofs follow from the definition of m.

***Theorem 1** (Invariance under Transpose).* (i) Let $A = [a_{ij}]$ and $C = [c_{ij}]$ be $n \times n$ judgment matrices. Then A^T and C^T are also judgment matrices, and

$$m(A^T,C^T) = m(A,C) .$$

(ii) Let $A = [a_{ij}]$ be an $n \times n$ judgment matrix, and suppose that $C = [c_{ij}]$ is the consistent matrix that is m-closest to A. Then C^T is the consistent matrix that is m-closest to A^T.

***Theorem 2** (Invariance under Change of Scale).* (i) Let $A = [a_{ij}]$ and $C = [c_{ij}]$ be $n \times n$ judgment matrices and $(w_1, w_2, \ldots, w_n)$ a ratio scale. Define

$$A' = [a_{ij} w_i / w_j]$$

$$C' = [c_{ij} w_i / w_j] .$$

Then A', C' are judgment matrices, and

$$m(A', C') = m(A, C) \ .$$

(ii) Let A, C, A', C' be as in (i), and suppose that C is the m-closest consistent matrix to A. Then C' is the m-closest consistent matrix to A'.

Recall that we seek a procedure for associating ratio scales to judgment matrices in such a way that the ratio scales capture the subjective information inherent in the corresponding matrices. Let A be an $n \times n$ judgment matrix. Let $C = [c_{ij}]$ be a consistent matrix that is m-closest to A, and suppose that $v = v_1, v_2, \ldots, v_n$ is a ratio scale for C; i.e.,

$$c_{ij} = \frac{v_i}{v_j} \ .$$

We choose $v_1, v_2, \ldots, v_n$ as the estimator of the ratio scale corresponding to A.

Under this association, *Theorem 1* guarantees that the scale $1/v_1, 1/v_2, \ldots, 1/v_n$ is the estimator of the scale corresponding to A^T. The appeal of this invariance arises in a natural way: Suppose a respondent put his estimate of u_i/u_j in the position r_{ji} instead of r_{ij}. (There appears to be nothing intrinsically right or wrong about recording the estimates this way.) In that case, the estimation procedure should return estimates $1/u_i$ instead of u_i. With this convention the estimated value of u_i should not depend on an artifact of the way the data are recorded. It follows from *Theorem 1* that the *GM* procedure has this invariance property. The *EV* procedure does not.[1]

Additionally, *Theorem 2* guarantees that our choice of ratio scale is invariant under a scale change in the judgment matrix. The eigenvector scale also fails this invariance test.

Theorem 3 guarantees that the geometric mean scale gives the m-closest consistent matrix to any judgment matrix.

Theorem 3. Let $A = [a_{ij}]$ be an $n \times n$ judgment matrix. Let $C = [c_{ij}]$ be the consistent matrix given by

$$c_{ij} = \frac{v_i}{v_j} \ ,$$

where v_i is the geometric mean of the elements of the ith row of A; i.e.,

$$v_i = \prod_{j=1}^{n} a_{ij}^{1/n}, \quad i = 1, 2, \ldots, n \ .$$

Then $m(A, C)$ is the minimal m-distance from A to any $n \times n$ consistent matrix.

Proof. For any consistent matrix $C = [c_{ij}]$, we can write

$$c_{ij} = \frac{w_i}{w_j} \ ,$$

where $w = w_1, w_2, \ldots, w_n$ is a ratio scale. Thus we seek a scale that minimizes the expression

$$\sum_{i=1}^{n} \sum_{j>i} [\ln a_{ij} - (\ln w_i - \ln w_j)]^2 \ .$$

[1]See Johnson, 1979, for an excellent discussion of this lack of symmetry and its relation to the choice between the right and left dominant eigenvector.

Because the estimating scale need be known only up to a scale factor, we may normalize by imposing the side condition

$$\prod_{i=1}^{n} w_j = 1 \ .$$

Let

$$Y_{ij} = \ln a_{ij}, \quad i,j = 1,2,\ldots,n \ ,$$
$$b_i = \ln w_i, \quad i = 1,2,\ldots,n \ .$$

Then the problem is to minimize

$$\sum_{i=1}^{n} \sum_{j>1}^{n} [y_{ij} - (b_i - b_j)]^2$$

under the side condition

$$\sum_{i=1}^{n} b_i = 0 \ .$$

Because

$$y_{ji} = -y_{ij} \qquad i,j = 1,2,\ldots,n$$

and

$$y_{ii} = 0 \ ,$$

this is equivalent to minimizing

$$S = \sum_{i=1}^{n} \sum_{j=1}^{n} [y_{ij} - (b_i - b_j)]^2$$

under the side condition

$$\sum_{i=1}^{n} b_i = 0 \ .$$

Now S is strictly convex in the differences $b_i - b_j$, and therefore strictly convex in the vector $\boldsymbol{b}$, so it has a unique minimum at the point where

$$\frac{\partial S}{\partial b_i} = 0 \text{ for } i = 1,2,\ldots,n \ .$$

Setting these partial derivatives equal to zero, for $k = 1,2,\ldots,n$

$$\frac{\partial S}{\partial b_k} = -2\sum(y_{kj} - b_k + b_j)$$

$$= -2\left(\sum_{j=1}^{n} y_{kj} - nb_k + \sum_{j=1}^{n} b_j\right) = 0 \ ,$$

and therefore, since $\sum_{j=1}^{n} b_j = 0$,

$$\sum_{j=1}^{n} y_{kj} = nb_k \ .$$

Thus S is minimized by

$$b_k = \frac{\sum_{j=1}^{n} y_{kj}}{n} \ ;$$

i.e.,

$$\ell n \ w_k = \frac{\sum_{j=1}^{n} \ell n \ a_{kj}}{n} \ , \quad k = 1,2, \ldots, n \ .$$

Consequently the m-distance from A to C is minimized by the vector v given by

$$v_k = \prod_{j=1}^{n} a_{kj}^{1/n} \ .$$

This completes the proof of *Theorem 3.*

Recall that if the matrix A is consistent, then the normalized geometric mean scale is equal to the normalized eigenvector scale. The two scales are always the same, regardless of consistency, if the dimension is less than or equal to three. To see this in the case $n = 3$, let

$$A = \begin{pmatrix} 1 & a & b \\ 1/a & 1 & c \\ 1/b & 1/c & 1 \end{pmatrix} \ .$$

Then the geometric mean vector for A,

$$v = \begin{pmatrix} (ab)^{1/3} \\ (c/a)^{1/3} \\ (1/bc)^{1/3} \end{pmatrix}$$

is an eigenvector for A corresponding to the eigenvalue

$$L = 1 + \left(\frac{ac}{b}\right)^{1/3} + \left(\frac{b}{ac}\right)^{1/3} \ .$$

Since L is of the form

$$1 + x + \frac{1}{x} \ ,$$

its value is no less than 3, the dimension of A (with equality in the consistent case). Therefore w is an eigenvector corresponding to the maximal eigenvalue for A. Hence in the case $n = 3$, the normalized geometric mean vector and the normalized eigenvector are the same. This result does not hold for inconsistent matrices with dimension greater than 3.

VII. THE GEOMETRIC MEAN VECTOR AND THE MAXIMUM LIKELIHOOD ESTIMATOR

We have shown that given an arbitrary judgment matrix A, the geometric mean vector gives rise to the m-closest consistent matrix to A. The problem of representing a judgment matrix by a ratio scale can also be cast in the framework of the general linear statistical regression model. We will show that under suitable assumptions on the distribution of errors in the expert's judgment, the geometric mean vector is the maximum likelihood estimator for the ratio scale corresponding to the judgment matrix.

Let $A = [a_{ij}]$ be an $n \times n$ judgment matrix. We assume that there is an underlying scale $w_1, w_2, \ldots, w_n$ whose ratios are perturbed (by errors of judgment) to give the elements of A, namely

$$a_{ij} = \frac{w_i}{w_j} e_{ij} \ ,$$

and thus

$$\ell n\ a_{ij} = \ell n\ w_i - \ell n\ w_j + \ell n\ e_{ij} \ , \tag{7.1}$$

$$i = 1,2,\ldots,n; j > i \ .$$

Regarding the distribution of the error term, the context of the Saaty-Vargas (1983) approach assumes a multiplicative model—if $u = u_1, u_2, \ldots, u_n$ is a scale for the entities $\{E_i\}$, then the value of E_i relative to E_j is given by u_i/u_j. Accordingly, we have assumed that errors are multiplicative. Further, this context assumes that if a_{ij} estimates u_i/u_j, then $1/a_{ji}$ is an equally good estimate of u_j/u_i; hence it is appropriate that the distribution of e_{ij} be reciprocal symmetric in the sense that

$$P(a < e_{ij} \leq b) = P\left(a < \frac{1}{e_{ij}} \leq b\right) \ .$$

Just as the normal distribution is a common model for additive errors, the log normal distribution, for similar reasons, is a common mathematical model for multiplicative errors. Additionally, it is reciprocal symmetric. We assume that model here. In Section VIII we evaluate the performance of the GM estimate under a very different error distribution.

We assume that the errors e_{ij} are independent and lognormally distributed with means 0 and variances σ^2. The substitution

$$Y = \begin{bmatrix} \ell n\ a_{1,1} \\ \ell n\ a_{1,2} \\ \vdots \\ \ell n\ a_{n-1,n} \end{bmatrix} \ , \quad B = \begin{bmatrix} \ell n\ w_1 \\ \vdots \\ \ell n\ w_n \end{bmatrix} \ , \quad E = \begin{bmatrix} \ell n\ e_{1,1} \\ \ell n\ e_{1,2} \\ \vdots \\ \ell n\ e_{n-1,n} \end{bmatrix}$$

gives the general linear equation

$$Y = XB + E,$$

where the matrix X has components $-1, 0, +1$ determined by Eq. (7.1). In this framework it is well known (Scheffe, 1959) that the maximum likelihood estimate for $B = [\ell n\ w_i]$ is the least-squares estimate given by

$$\hat{b}_i = \frac{1}{n} \sum_{j=1}^{n} \ell n\ a_{ij} \ ,$$

and that the estimate has all of the usual desirable properties of least-squares estimates under the general linear hypothesis (unbiased, minimum variance, etc.). Taking exponentials, we find that the maximum likelihood estimate of w_i is given by:

$$\hat{w}_i = \exp(\hat{b}_i) = \prod_{j=1}^{n} a_{ij}^{1/n} \ .$$

(The same estimate is derived above from the metric m on the space of judgment matrices.)

The procedure outlined above can be modified to solve more general estimation problems. For example, suppose that instead of a single comparison for each pair of objects E_i and E_j, there are n_{ij} comparisons, a_{ijk}, $k = 1, \ldots, n_{ij}$ where n_{ij} may be zero (reflecting missing data) or greater than one (reflecting multiple comparisons, say by different judges). The problem is then to find a vector w that minimizes the sum of squares:

$$S = \sum_{i=1}^{n} \sum_{j>1} \sum_{k=1}^{n_{ij}} [\ell n\ a_{ijk} - (\ell n\ w_i - \ell n\ w_j)]^2 \ .$$

This generalization does not yield a simple closed-form solution such as the geometric mean vector, but in practice S can be minimized and w determined using standard least-squares regression packages.

Regardless of the values of n_{ij}, the geometric mean estimation procedure leads to a natural measure of consistency for judgment matrices that is well grounded in statistical theory and can be used in hypothesis testing. Let s^2 be the residual mean square

$$s^2 = \frac{S}{d.f.} \ ,$$

where $d.f.$ is the number of independent observations minus the number of linearly independent parameters.

Note that if $n_{ij} = 1$, then

$$d.f. = \frac{n(n-1)}{2} - (n-1) = \frac{(n-1)(n-2)}{2} \ .$$

Then s^2 is an unbiased estimator of σ^2 (the variance of the perturbations) and hence is a natural measure of consistency of A.

Recall that if $n_{ij} = 1$, then S can be viewed as the squared distance from A to the m-closest consistent matrix. Therefore S is zero when A is consistent, is close to zero when A is close to consistent, and is far from zero when A is far from consistent. Moreover, because S depends entirely on ratios, it is invariant under scale changes and transposes in the sense of *Theorems 1 and 2.*

VIII. EMPIRICAL COMPARISON OF THE TWO METHODS: A MONTE CARLO STUDY

The geometric mean vector gives an estimate for ratio scales based on pairwise comparisons that is easy to calculate, satisfies the theoretical requirements of invariance under scale change and transpose, and is well grounded in statistical theory. It follows from Section VI that under certain assumptions on the distribution of the perturbations, the logarithms of the geometric mean vector components are the minimum variance unbiased estimators of the logarithms of the underlying ratio scale factors. Thus, if errors have a lognormal distribution, the *GM* procedure provides the best estimate of the underlying scale and in particular is preferable to the *EV* procedure with respect to the metric given above.

A good procedure should be robust in the sense that it yields good performance when measured by other relevant metrics and different choices for the distribution of errors. Although the robustness of the General Linear Model has been widely treated in the literature, we look here at several very specific comparisons of the performance of the *GM* and *EV* procedures.

In this section we give the results of Monte Carlo trials that compare the *EV* and the *GM* procedures under different choices of error distributions and different choices of metrics. These studies imply that the *EV* is a better estimator (in the sense of metrics considered here) under both choices of distributions of errors. The metrics used include measures of rank preservation, where the *EV* has been purported to be better than the *GM* procedure (see appendix).

$$a_{ij} = \frac{u_i}{u_j} \ .$$

Throughout this section the u_i will be equal to $i \times R$, where the constant R is taken so that the u_i sum to one.

Let $\{e_{ij}, i,j = 1, \ldots, n, j > i\}$ be a collection of positive independent random variables drawn from a suitable population; we construct a perturbed matrix D with elements

$$d_{ij} = \begin{cases} \dfrac{u_i}{u_j} e_{ij} & \text{for } i = 1,2,\ldots,n,\ j > i \\ \dfrac{1}{d_{ji}} & \text{for } i = 1,2,\ldots,n,\ j < i \\ 1 & \text{for } i = 1,2,\ldots,n,\ j = i \end{cases} .$$

D is a judgment matrix, but because of the perturbations it is typically not consistent. Any estimate of an underlying ratio scale for D should give rise to a consistent matrix that is in some sense close to A. Two quite different metrics for closeness[1] are used in the Monte Carlo

[1]It has been mentioned in the literature (Saaty and Vargas, 1984, and others), that Fichtner (1984) has developed a metric m with the property that the m-closest consistent matrix to a subjective judgment matrix is the consistent matrix corresponding to the maximal eigenvector. Typically a metric is of interest because it provides an intuitive understanding of the topology of a space—i.e., it tells you what points are close to (neighborhoods of) other points, hence what sequences converge, and to what points. Unfortunately, Fichtner's complicated metric yields the discrete topology, a very pathological example wherein every point is an open set, the only sequences that converge are those whose terms are all the same from some point on, and even Saaty and Vargas's (1984) normalized limit of A^n does not exist. It is the same topology given by the metric that calls the distance from two points 1 if the points are not equal and 0 if they are.

study described here: the sum of squares of errors (*SSE*)

$$SSE = \sum_{i=1}^{n} (\hat{u}_i - u_i) \ ,$$

and the sum of squares of the errors of logarithms:

$$SSEL = \sum_{i=1}^{n} (\ell n \ \hat{u}_i - \ell n \ u_i) \ ,$$

where

$$u_i, i = 1,3, \ldots, n$$

is the actual normalized ratio scale, and

$$\hat{u}_i, i = 1,2, \ldots, n$$

is the estimated ratio scale. In addition we have looked at several measures of order preservation, where the *EV* procedure is said to be preferable to the *GM* procedure.

The choice of population from which the multiplicatively perturbing random variables e_{ij} are drawn should reflect that in a judgment matrix the distribution of errors should be reciprocal symmetric as mentioned above:

$$P(a < e_{ij} \leq b) = P\left(a < \frac{1}{e_{ij}} \leq b\right) \ .$$

We have used two convenient and very dissimilar distributions satisfying this property: a log-normal distribution whose underlying normal distribution has mean zero, and a distribution obtained from the ratio of two independent, uniformly distributed (on (c,d) where $0 < c < d$) random variables. We consider perturbations e_{ij} drawn from populations with distributions of both of these types.

We computed the sums of squares of errors and sums of squares of errors of logarithms for individual trials and then totaled them over the trials for both the geometric mean vector and the eigenvector estimates.

In a comparison of two estimation techniques A and B, it is common to look at the average or sum of performance measures over some large number of trials or, when possible, the expected performance with respect to some distribution of errors. Because most researchers are not faced with the problem of routinely analyzing thousands of subjective judgment matrices, these results may be misleading: Although A may be preferable to B in the sense of expected errors, B may be preferable to A 90 percent of the time, the seeming contradiction being explained by B occasionally giving large outliers. When an outlier is suspected, the researcher may reject that method on that trial and look at other methods. Because of these concerns we have scored each of the Monte Carlo trials and computed the frequency with which the geometric mean outperforms the eigenvector. The perverse situation mentioned above does not seem to occur here. Within experimental error, the geometric mean performs as well as, or better than, the eigenvector in the sum of errors and outperforms it in 50 percent or more of the trials.

Comparison with Sum of Squared Log Error and Lognormal Errors

Table 2 gives the results of the trial that will theoretically favor the geometric mean vector: lognormal errors and the metric given by the sum of the squared errors of logarithms. The entries in the first column describe the parameters of the Monte Carlo run. The first number is the dimension of the scale, the second number is the variance of the logarithm of the error term e_{ij}. The next term is the average "consistency ratio" for the trial—that is, the difference between the maximal eigenvalue and the dimension, normalized by the dimension minus 1. We are uncertain about what the consistency ratio really means, but the eigenvector is said to do well when the consistency ratio is less than .10. Accordingly we have concentrated on runs in this realm. The next number gives the number of trials. In the cases with very small log perturbation error we used 5000 trials, in the cases where the trend seemed clear we used 1000 trials.[2]

The eigenvector outperforms the geometric mean in every case, in terms of both sum of errors and the percentage of trials where it was closer to the underlying consistent matrix.

Table 2

RATIO SCALES PERTURBED BY LOGNORMAL ERRORS, SUM OF SQUARED ERRORS OF LOGARITHMS, MONTE CARLO COMPARISON OF GEOMETRIC MEAN VECTOR AND EIGENVECTOR

Dimension, Variance, $(L-n)/(n-1)$, Number of Trials	Sum of Squared Error of Logs, Geometric Mean	Sum of Squared Error of Logs, Eigenvector	Percentage of Trials in Which Geometric Mean Is Better in Squared Error of Logs
5/.01/.003/5000	42.4	42.5	51
5/.04/.012/5000	170.8	171.7	52
5/.09/.028/5000	377.6	382.1	53
5/.16/.051/1000	135.2	138.4	57
5/.25/.080/1000	210.4	218.8	60
5/.49/.165/1000	427.0	446.6	56
5/1.0/.349/1000	863.3	947.7	61
7/.01/.005/5000	45.2	45.3	53
7/.04/.015/5000	178.9	180.6	55
7/.09/.033/5000	403.5	410.9	55
7/.16/.059/1000	150.9	154.5	56
7/.25/.095/1000	222.7	235.1	59
7/.49/.208/1000	440.6	485.2	61
7/1.0/.445/1000	928.2	1102	67
10/.01/.040/5000	46.4	46.5	53
10/.04/.016/5000	185.4	187.6	56
10/.09/.036/5000	421.9	433.4	60
10/.16/.067/1000	147.7	156.2	66
10/.25/.108/1000	234.2	251.6	65
10/.49/.223/1000	456.2	524.2	69
10/1.0/.516/1000	946.1	1268	78

[2]These runs were made in Fortran on a Compaq personal computer. The principal author will gladly supply the programs to the reader wishing to evaluate his own scenarios.

The preference for the eigenvector is weakest when the dimension or the error is small, as would be expected because the methods give the same answer when the dimension is 3 or less, or when the matrix is consistent.

Comparison with Sum of Squared Log Error and Ratio of Uniform Errors

Table 3 presents the same information for the case of perturbations by ratios of uniform random variables. The variance given is (as above) the variance of the logarithm of the perturbation term. The consistency ratio is determined empirically and may differ in the third decimal place. Even with these very different error terms, the geometric mean vector is preferable in every case and in both measures. Again, the degree of preference is small when the methods can be expected to give answers that are very close and becomes much stronger as the dimension or the variance of errors grows.

Table 3

RATIO SCALES PERTURBED BY RATIOS OF UNIFORM RANDOM VARIABLES, SUM OF SQUARED ERRORS OF LOGARITHMS, MONTE CARLO COMPARISON OF GEOMETRIC MEAN VECTOR AND EIGENVECTOR

Dimension, Variance, $(L-n)/(n-1)$, Number of Trials	Sum of Squared Error of Logs, Geometric Mean	Sum of Squared Error of Logs, Eigenvector	Percentage of Trials in Which Geometric Mean Is Better in Squared Error of Logs
5/.01/.003/5000	42.3	42.3	50
5/.04/.012/5000	167.8	168.1	51
5/.09/.027/5000	385.5	385.4	50
5/.16/.050/1000	134.4	136.0	51
5/.25/.081/1000	214.3	219.8	56
5/.49/.161/1000	425.1	442.2	55
5/1.0/.360/1000	857.4	969.7	66
7/.01/.004/5000	44.3	44.3	51
7/.04/.014/5000	177.5	177.9	52
7/.09/.033/5000	406.2	410.5	53
7/.16/.059/1000	145.0	146.0	51
7/.25/.094/1000	225.8	229.9	55
7/.49/.195/1000	453.3	492.6	65
7/1.0/.451/1000	906.1	1119	72
10/.01/.004/5000	46.2	46.2	52
10/.04/.016/5000	185.6	186.4	53
10/.09/.037/5000	415.8	420.7	55
10/.16/.066/1000	148.8	151.3	55
10/.25/.106/1000	230.3	239.1	61
10/.49/.219/1000	460.2	505.9	66
10/1.0/.519/1000	953.3	1274	82

Comparison with Squared Errors

Tables 4 and 5 consider the sum of squared error metric instead of the sum squared error of logarithms. Lognormal errors have been used in Table 4 and the ratio of uniform errors in Table 5. In this metric, which differs substantially from the metric used to justify the geometric mean, the geometric mean still outperforms the eigenvector in both measures in both tables in every case except one. In the "errant" case the eigenvector does better by "1" in the least significant digit. Given the trend of the results, we conclude that this aberration is the result of experimental error. The pattern mentioned above repeats itself: The big differences occur when the dimension or the variance of the errors becomes large.

Table 4

RATIO SCALES PERTURBED BY LOGNORMAL ERRORS, SUM OF SQUARED ERRORS, MONTE CARLO COMPARISON OF GEOMETRIC MEAN VECTOR AND EIGENVECTOR

Dimension, Variance, $(L-n)/(n-1)$, Number of Trials	Sum of Squared Error, Geometric Mean	Sum of Squared Error, Eigenvector	Percentage of Trials in Which Geometric Mean is Least Squares Better
5/.01/.003/5000	1.70	1.70	51
5/.04/.012/5000	6.81	6.84	52
5/.09/.028/5000	15.0	15.1	53
5/.16/.051/1000	3.13	3.15	52
5/.25/.080/1000	8.15	8.50	56
5/.49/.165/1000	16.1	16.9	54
5/1.0/.349/1000	31.3	34.9	58
7/.01/.005/5000	1.00	1.01	52
7/.04/.015/5000	3.92	3.96	53
7/.09/.033/5000	8.94	9.11	54
7/.16/.059/1000	3.33	3.43	55
7/.25/.095/1000	4.81	5.08	57
7/.49/.203/1000	9.77	10.7	58
7/1.0/.445/1000	19.4	23.3	61
10/.01/.040/5000	.538	.539	52
10/.04/.016/5000	2.15	2.16	53
10/.09/.036/5000	4.90	5.04	56
10/.16/.067/1000	1.73	1.83	59
10/.25/.108/1000	2.71	2.90	61
10/.49/.223/1000	5.15	5.88	64
10/1.0/.516/1000	10.5	14.1	71

Table 5

RATIO SCALES PERTURBED BY RATIOS OF UNIFORM RANDOM VARIABLES, SUM OF SQUARED ERROR, MONTE CARLO COMPARISON OF GEOMETRIC MEAN VECTOR AND EIGENVECTOR

Dimension, Variance, $(L - n)/(n - 1)$, Number of Trials	Sum of Squared Error, Geometric Mean	Sum of Squared Error, Eigenvector	Percentage of Trials in Which Geometric Mean is Least Squares Better
5/.01/.003/5000	1.71	1.71	50
5/.04/.012/5000	6.73	6.75	51
5/.09/.027/5000	15.7	15.6	50
5/.16/.050/1000	5.38	5.39	51
5/.25/.081/1000	8.40	8.55	53
5/.49/.161/1000	16.1	16.6	55
5/1.0/.360/1000	31.2	35.1	62
7/.01/.004/5000	.996	.996	50
7/.04/.014/5000	3.96	3.96	50
7/.09/.033/5000	9.07	9.13	53
7/.16/.059/1000	3.13	3.14	51
7/.25/.094/1000	4.92	4.98	52
7/.49/.195/1000	9.80	10.6	61
7/1.0/.451/1000	19.1	24.0	70
10/.01/.004/5000	.531	.531	50
10/.04/.016/5000	2.14	2.15	51
10/.09/.037/5000	4.76	4.79	53
10/.16/.066/1000	1.73	1.75	52
10/.25/.106/1000	2.67	2.75	56
10/.49/.219/1000	5.29	5.71	61
10/1.0/.519/1000	10.7	14.4	74

IX. EXAMPLE

In this section we discuss the use of the eigenvector and the geometric mean vector in a specific subjective judgment situation. Estimates of underlying utility vectors and consistency values derived from the two methods are compared.

Consider the automobile selection problem introduced in Sections II and V. The hierarchical structure for this problem was given above in Fig. 1.

Subjective judgment data for this example were obtained from one prospective buyer. The buyer made pairwise comparisons reflecting his perceptions of the relative importance of the attributes of status, cost, economy, and size in selecting an automobile. Judgments were made based on the subjective judgment scale developed by Saaty (1977b). The buyer made comparisons in such a way that the resulting pairwise comparison matrix would be reciprocal symmetric (i.e., a judgment matrix). The resulting judgment matrix A is:

	Status	Cost	Economy	Size
Status	1	1/5	1/5	1/2
Cost	5	1	1	1/3
Economy	5	1	1	1/2
Size	2	3	2	1

$= A$

Next, we constructed judgment matrices from the buyer's pairwise comparisons of the five types of automobiles relative to each of the two attributes of status and size:

Subjective Comparison Relative to Status

	H	T	M	D	C
H	1	2	1/9	1/6	1/7
T	1/2	1	1/9	1/6	1/7
M	9	9	1	6	4
D	6	6	1/6	1	1/7
C	7	7	1/4	7	1

$= B$

Subjective Comparison Relative to Size

	H	T	M	D	C
H	1	1/2	7	3	8
T	2	1	8	4	9
M	1/7	1/8	1	1/3	3
D	1/3	1/4	3	1	4
C	1/8	1/9	1/3	1/4	1

$= C$

We calculated the eigenvector and the geometric mean vector separately for each of the three subjective judgment matrices above. Resulting scale estimates for the two techniques are given in Table 6.

Notice that the scales determined by the two methods are very close in value. This is as expected from theoretical considerations, because the two methods give the same results for consistent matrices and should agree closely for nearly consistent ones.

The results in Table 6 indicate that the prospective buyer considers automobile size to be considerably more important than status, cost, or economy. He ranks T as having the best size of any of the cars under consideration.

Because exact values were available for cost and economy of the five automobile types, it was not necessary to compute scales for them from pairwise comparisons. The normalized scale values for the automobiles relative to these two attributes were determined to be:

	Cost Ratings	Economy Ratings
H	.2703	.3357
T	.4196	.2929
M	.0554	.1071
D	.1571	.1500
C	.0976	.1143

Table 6

COMPARISON OF SCALE ESTIMATES FOR THE AUTOMOBILE EXAMPLE

	Matrix					Normalized Geometric Mean	Consistency s^2	Normalized Eigenvector	Consistency μ
			Matrix A						
Status	1	1/5	1/5	1/2		.0812	.5680	.0885	.1513
Cost		1	1	1/3		.2453		.2425	
Economy			1	1/2		.2715		.2579	
Size				1		.4019		.4112	
			Matrix B Status						
H	1	2	1/9	1/6	1/7	.0409	.5550	.0378	.1770
T		1	1/9	1/6	1/7	.0310		.0294	
M			1	6	4	.5307		.5239	
D				1	1/7	.1132		.1131	
C					1	.2842		.2958	
			Matrix C Size						
H	1	1/2	7	3	9	.3152	.1150	.3143	.0417
T		1	8	4	9	.4632		.4630	
M			1	1/3	3	.0581		.0596	
D				1	4	.1299		.1288	
C					1	.0336		.0343	

From each method, we now have scale estimates for the five automobile types relative to each of the four automobile attributes, as well as a scale of importance of the four attributes. Scale estimates for the five cars relative to each attribute are used as columns of a 5 × 4 matrix. This matrix is multiplied by the 4-dimensional vector of importance of attributes. The resulting 5-dimensional vector reflects the prospective buyer's ranking of the five automobiles. The calculations for both methods are carried out in Table 7.

Table 7

DETERMINATION OF RANKINGS OF FIVE AUTOMOBILES

Make of Car	Status	Cost	Economy	Size		Criterion Scale		Final Scale
Geometric Mean Vector Method								
H	.0409	.2703	.3357	.3152		.0812		.2874
T	.0310	.4196	.2929	.4632		.2453		.3711
M	.5307	.0554	.1071	.0581	×	.2715	=	.1091
D	.1132	.1571	.1500	.1299		.4019		.1407
C	.2842	.0976	.1143	.0336				.0916
Eigenvector Method								
H	.0378	.2703	.3357	.3143		.0885		.2847
T	.0294	.4196	.2929	.4630		.2425		.3703
M	.5239	.0554	.1071	.0596	×	.2579	=	.1120
D	.1131	.1571	.1500	.1288		.4112		.1398
C	.2958	.0976	.1143	.0343				.0934

Appendix

COMPARISON OF RANK PRESERVATION ABILITIES

We believe the issue of rank preservation is a red herring: The *EV* procedure has been recommended for analyzing hierarchical structures and quantifying judgments. A procedure designed to do well in a cardinal estimation problem ought not be judged on the basis of its capabilities in ordinal or rank estimation problems. If the researcher has an ordinal estimation problem it would be better and easier to ask the respondent to rank the entities in order of preference.

Further, the oft-mentioned ability of the *EV* to do well in this context seems to be based (Saaty and Vargas, 1984) on the following argument:

> There is a natural way to derive the rank order of a set of alternatives from a pairwise comparison matrix A. The rank order of each alternative is the relative proportion of its dominance over the other alternatives. *This is obtained by adding the elements in each row in A and dividing by the total over all rows.* (Emphasis added.)

Everything that has been written in the present context about the analysis of subjective judgment matrices has dealt with multiplicative problems. The estimates are of ratios, not of sums or differences. In this context determining rank order by adding elements seems unnatural. The natural ranking inherent in a subjective judgment matrix for a multiplicative problem is that obtained by the product of the row entries, not the sum. The *GM* procedure gives precisely that ranking, and it outperforms the *EV* procedure in this measure also.

Although we believed the issue was not worth treating, referees objected to the omission of rank preservation properties in the comparisons of the two methods, hence its inclusion here. In Tables A.1 and A.2 we have measured the ability of the *GM* and *EV* procedures to preserve rank. The scenarios are as above: Table A.1 uses lognormal errors and Table A.2 ratios of uniform random variables. Recall that the underlying scales u_i are proportional to i and scaled so that they sum to one.

Rank is typically expressed as a vector; we know of no single overwhelmingly natural scalar measure of rank error. We have used two very different measures, both simple, but lacking simple descriptions. To fix ideas, let us assume that the dimension of the scale is 5. We represent the ranking in the underlying scale by 1,2,3,4,5. If we estimate this scale by a vector v that satisfies

$$v_1 < v_2 < v_5 < v_3 < v_4 \ ,$$

then we represent the ranking of v by 1,2,5,3,4. If v preserves the ranking inherent in u, its ranking will be the same as that of u.

We have called the measures of rank preservation used below the sum of rank reversals and the sum of weighted rank reversals. The sum of rank reversals is just the number of times we count a reversal in ordering as we go from left to right ("1" in the above example, because $5 > 3$). The weighted rank reversal measures the degree of displacement from the rank of the underlying scale. In the above example it is given by

$$(1 - 1)^2 + (2 - 2)^2 + (3 - 5)^2 + (4 - 3)^2 + (5 - 4)^2 = 6.$$

Table A.1

RATIO SCALES PERTURBED BY LOGNORMAL ERRORS, RANK PRESERVATION ERROR, MONTE CARLO COMPARISON OF GEOMETRIC MEAN VECTOR AND EIGENVECTOR

Dimension, Variance, $(L-n)/(n-1)$, Number of Trials	Sum of Rank Reversal Errors, Geometric Mean	Sum of Rank Reversal Errors, Eigenvector	Sum of Weighted Rank Reversal Errors, Geometric Mean	Sum of Weighted Rank Reversal Errors, Eigenvector
5/.01/.003/5000	2	2	4	4
5/.04/.012/5000	258	260	516	520
5/.09/.028/5000	963	971	1964	1980
5/.16/.051/1000	210	209	426	422
5/.25/.080/1000	529	538	1160	1190
5/.49/.165/1000	825	847	2002	2060
5/1.0/.349/1000	1081	1141	2874	3036
7/.01/.005/5000	14	14	28	28
7/.04/.015/5000	683	694	1372	1394
7/.09/.033/5000	2084	2143	4384	4510
7/.16/.059/1000	712	739	1588	1650
7/.25/.096/1000	911	962	2178	2298
7/.49/.203/1000	1332	1375	3548	3694
7/1.0/.445/1000	1775	1826	5322	5642
10/.01/.040/5000	74	77	148	154
10/.04/.016/5000	1809	1824	3684	3730
10/.09/.036/5000	4350	4446	9590	9832
10/.16/.067/1000	1317	1367	3174	3330
10/.25/.108/1000	1659	1710	4334	4534
10/.49/.223/1000	2189	2282	6460	6948
10/1.0/.516/1000	2667	2852	9228	10266

In Tables A.1 and A.2 we have given the sum of rank reversal errors and the sum of weighted rank reversal errors. (In the tables in this report we started each new Monte Carlo process with a new random number seed to avoid repeatedly using a sequence of random numbers that may favor one procedure or the other.)

In Table A.1, with lognormal perturbations, in only two out of 42 comparisons does the eigenvector outperform the geometric mean. In four other comparisons the methods tied. In the remaining 36 pairs, the *GM* procedure outperforms the *EV* procedure, and the differences in performance become more pronounced as the dimension and underlying variance increase.

In Table A.2, where the perturbations are ratios of uniform random variables, the eigenvector also does poorly, although a little better than above. Again, the relative performance of the geometric mean improves as the dimension and the variance of the errors increase. Even in these trials, which are dominated by the low variance and consistency ratios that are said to favor the eigenvector, the few occasions where the eigenvector outperforms the geometric mean seem to be the result of random aberrations, not of any inherent qualities that are pertinent to estimating ratio scales.

Table A.2

RATIO SCALES PERTURBED BY RATIOS OF UNIFORM RANDOM VARIABLES, RANK PRESERVATION ERROR, MONTE CARLO COMPARISON OF GEOMETRIC MEAN VECTOR AND EIGENVECTOR

Dimension, Variance, $(L-n)/(n-1)$, Number of Trials	Sum of Rank Reversal Errors, Geometric Mean	Sum of Rank Reversal Errors, Eigenvector	Sum of Weighted Rank Reversal Errors, Geometric Mean	Sum of Weighted Rank Reversal Errors, Eigenvector
5/.01/.003/5000	0	0	0	0
5/.04/.012/5000	252	248	504	496
5/.09/.027/5000	1069	1074	2164	2172
5/.16/.050/1000	343	340	724	716
5/.25/.081/1000	533	539	1144	1178
5/.49/.161/1000	779	824	1890	2002
5/1.0/.360/1000	1077	1125	2888	3094
7/.01/.004/5000	4	4	8	8
7/.04/.014/5000	685	702	1374	1410
7/.09/.033/5000	2154	2176	4530	4576
7/.16/.059/1000	715	699	1558	1548
7/.25/.094/1000	953	938	2286	2288
7/.49/.195/1000	1322	1362	3532	3706
7/1.0/.451/1000	1692	1794	5116	5758
10/.01/.004/5000	72	70	144	140
10/.04/.016/5000	1760	1756	3594	3598
10/.09/.037/5000	4168	4228	9220	9354
10/.16/.066/1000	1295	1284	3078	3112
10/.25/.106/1000	1632	1661	4264	4400
10/.49/.219/1000	2161	2212	6434	6844
10/1.0/.519/1000	2714	2917	9406	10602

BIBLIOGRAPHY

Aczel, J., and T. L. Saaty, "Procedures for Synthesizing Ratio Judgments," *Journal of Mathematical Psychology*, Vol. 27, No. 1, March 1982, pp. 93–102.

Alexander, J. M., and T. L. Saaty, "The Forward and Backward Processes of Conflict Analysis," *Behavioral Science*, Vol. 22, 1977, pp. 87–98.

Ayres, R. U., *Technological Forecasting and Long-range Planning*, McGraw-Hill, New York, 1969.

Churchman, C. W., and R. L. Ackoff, "An Approximate Measure of Value," *Journal of the Operations Research Society of America*, Vol. 2, No. 2, May 1954, pp. 172–187.

David, H. A., *The Method of Paired Comparisons*, Hafner Press, Macmillan, New York, 1963.

Davidson, R. R., and P. H. Farquhar, "A Bibliography on the Method of Paired Comparisons," *Biometrics*, Vol. 32, June 1976, pp. 241–252.

Fichtner, J. J., *Some Thoughts About the Mathematics of the Analytic Hierarchy Process*, Hochschule der Bundeswehr, Munchen, June 1984.

Franklin, J. N., *Matrix Theory*, Prentice-Hall, Englewood Cliffs, New Jersey, 1968.

Gordon, T. J., and O. Helmer, "Reporting on a Long-range Forecasting Study," The Rand Corporation, P-2982, September 1964.

Johnson, C. R., "Right-left Asymmetry in an Eigenvector Ranking Procedure," *Journal of Mathematical Psychology*, Vol. 79, 1977, pp. 61–64.

Keeney, R. L., "A Decision Analysis with Multiple Objectives: The Mexico City Airport," *Bell Journal of Economics and Management Science*, Vol. 4, 1973, pp. 101–117.

Keeney, R. L., and K. Nair, "Decision Analysis for Siting of Nuclear Plants: The Relevance of Multiattribute Utility Theory," Technical Report No. 96, Operations Research Center, MIT, 1974.

Keeney, R. L., and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, John Wiley and Sons, New York, 1976.

Kelly, C. W., III, "Decision Analysis and Decision Aids," *Proceedings of the Workshop on Decision Information for Tactical Command and Control*, Airlie, Virginia, September 22–25, 1976, pp. 287–301.

Linstone, H. A., and M. Turoff, *The Delphi Method—Techniques and Applications*, Addison-Wesley Publishing Company, Inc., Reading, Mass., 1975.

MacCrimmon, K. R., "Improving the System Design and Evaluation Process by the Use of Trade-off Information: An Application to Northeast Corridor Transportation Planning," The Rand Corporation, RM-5877-DOT, April 1969.

North, H. Q., and D. L. Pyke, "Technology, the Chicken—Corporate Goals, the Egg," *1st Annual Technology and Management Conference*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1968.

Press, S. J., "Qualitative Controlled Feedback for Forming Group Judgments and Making Decisions," The Rand Corporation, P-6290, January 1979.

Raiffa, H., "Preferences for Multiattributed Alternatives," The Rand Corporation, RM-5868-DOT/RC, April 1969.

Saaty, T. L., "Operation Gaming for Energy-Policy Analysis," *Energy Policy*, Vol. 5, No. 1, 1977, pp. 63–75a.

Saaty, T. L., "A Scaling Method for Priorities in Hierarchical Structures," *Journal of Mathematical Psychology*, Vol. 15, 1977b, pp. 234–281.

Saaty, T. L., "Scenarios and Priorities in Transport Planning: Application to the Sudan," *Transportation Research*, Vol. 11, 1977c, pp. 343–350.

Saaty, T. L., "The Sudan Transport Study," *Interfaces*, Vol. 8, No. 1, Part 2, November 1977d, pp. 37–57.

Saaty, T. L., and J. P. Bennett, "A Theory of Analytical Hierarchies Applied to Political Candidacy," *Behavioral Science*, Vol. 22, 1977, pp. 237–245.

Saaty, T. L., and M. W. Khouja, "A Measure of World Influence," *Journal of Peace Science*, Vol. 2, No. 1, Spring 1976, pp. 31–48.

Saaty, T. L., and P. C. Rogers, "Higher Education in the United States (1985–2000)," *Socio-Economic Planning Sciences*, Vol. 10, No. 6, 1976, pp. 251–263.

Saaty, T. L., and P. C. Rogers, "Scenario Construction Using a Hierarchical Framework with Eigenvector Weighting," *Socio-Economic Planning Sciences*, Vol. 10, Pergamon Press, Great Britain, 1976, pp. 251–263.

Saaty, T. L., and L. G. Vargas, "Inconsistency and Rank Preservation," *Journal of Mathematical Psychology*, Vol. 28, No. 2, March 1983, pp. 205–214.

Scheffe, H., *The Analysis of Variance*, John Wiley and Sons, New York, 1959.

Sigford, J. V., and R. H. Parvin, "Project PATTERN: A Methodology for Determining Relevance in Complex Decision-Making," *IEEE Transactions on Engineering Management*, Vol. EM-12, March 1965.

Torgerson, W. S., *Theory and Methods of Scaling*, John Wiley and Sons, New York, 1958.

Veit, C. T., and M. Callero, *Subjective Transfer Function Approach to Complex System Analysis*, The Rand Corporation, R-2719-AF, March 1981.

Veit, C. T., M. Callero, and B. J. Rose, *Demonstration of the Subjective Transfer Function Approach Applied to Air-Force-Wide Mission Area Analysis*, The Rand Corporation, N-1831-AF, February 1982.

Veit, C. T., M. Callero, and B. J. Rose, *Introduction to the Subjective Transfer Function Approach to Analyzing Problems*, The Rand Corporation, R-3021-AF, July 1984.

von Neumann, J., and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, New Jersey, 2d ed., 1947.

Wells, H. A., "Weapon System Planner's Guide," *IEEE Transactions on Engineering Management*, Vol. EM-14, March 1967, pp. 14–16.

END

FILMED

9-85

DTIC